

# SASI

Sichere und Akzentrobuste Sprachassistenz für Intensivstationen  
(Secure & Accent-Robust Voice Assistance for Intensive Care Units)

1.8M

healthcare positions unfilled  
in Germany by 2035

15%

of practicing physicians in  
Germany are non-German  
nationals

~1h

per shift saved through voice-  
assisted documentation

30%

efficiency gain in  
documentation tasks

## Section 1

### The challenge

#### Post-bypass Intensive Care Unit (ICU):

Patient Blood pressure (BP) dropping. Nurse needs to administer vasopressor immediately.

#### ICU nurse speaks

Foreign-trained ·  
accented German

Voice command  
"Noradrenalin null  
Komma eins  
Mikrogramm pro  
Kilogramm pro  
Minute"

Intended dose:  
0.1 mcg/kg/min

#### Automatic Speech Recognition (ASR) system transcribes

Standard ASR · no accent adaptation

Said:  
0.1 mcg/kg/min

Transcribed as  
1.0 mcg/kg/min

"null Komma eins" the German  
decimal phrasing was not  
recognized. The system heard a  
ten times higher dose.

#### No plausibility check

Input accepted · no alert  
raised

Patient record KIS 03:17  
BP check 78/42 mmHg  
Noradrenalin logged 1.0  
mcg/kg/min  
Pump adjusted 03:19

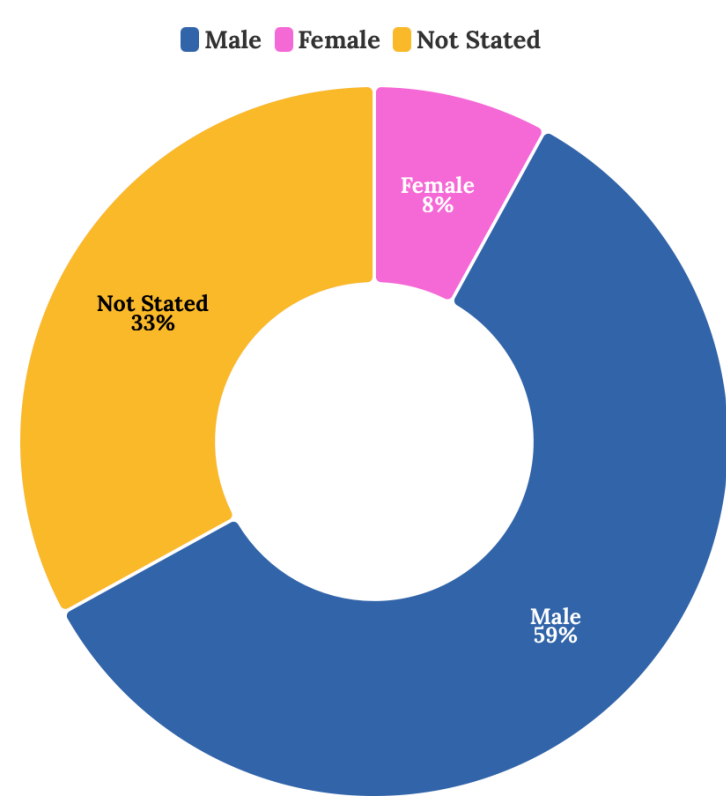
Error noticed only when  
patient deteriorated  
minutes later. No system  
flagged the implausible  
dose.

In intensive care, a misrecognized command is not a minor bug, it is a patient safety event. Current ASR systems have no mechanism to catch this.

## Section 2

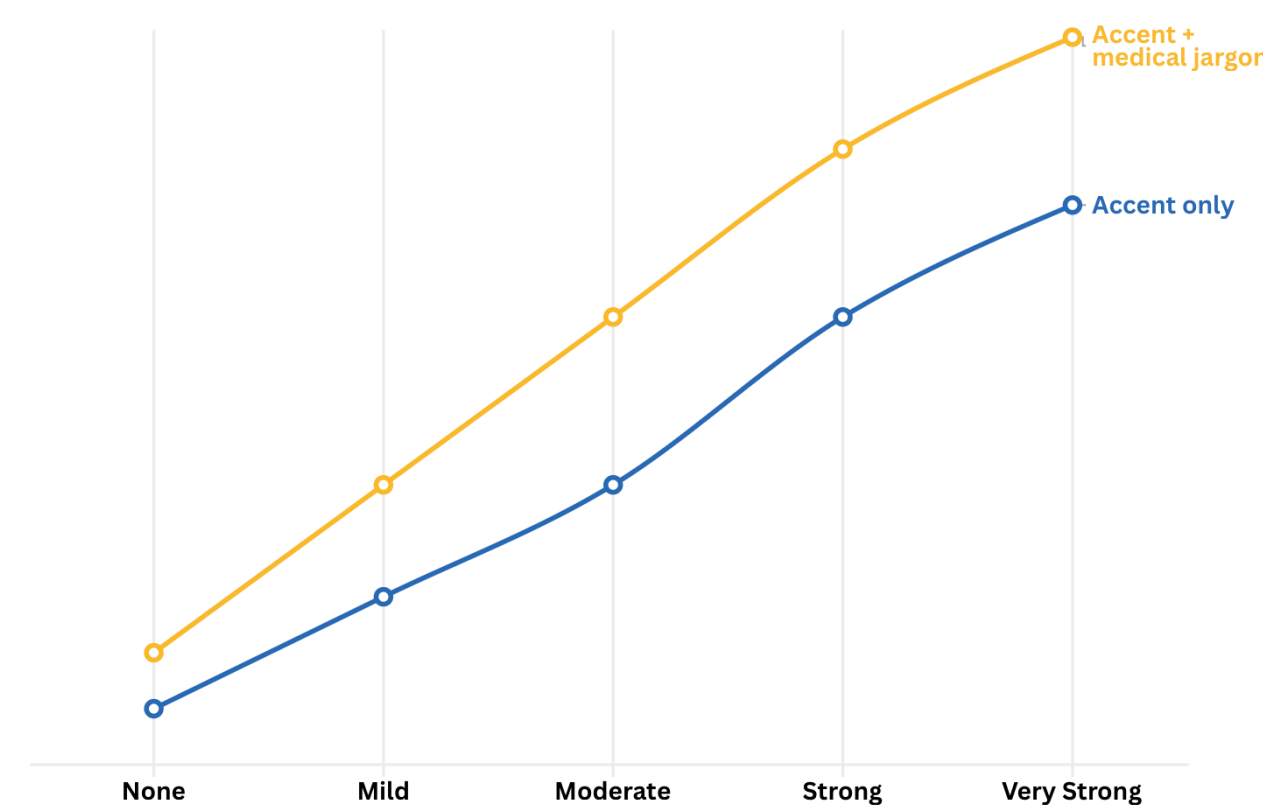
### Limitations of current technology

Training data gender distribution  
Mozilla Common Voice dataset



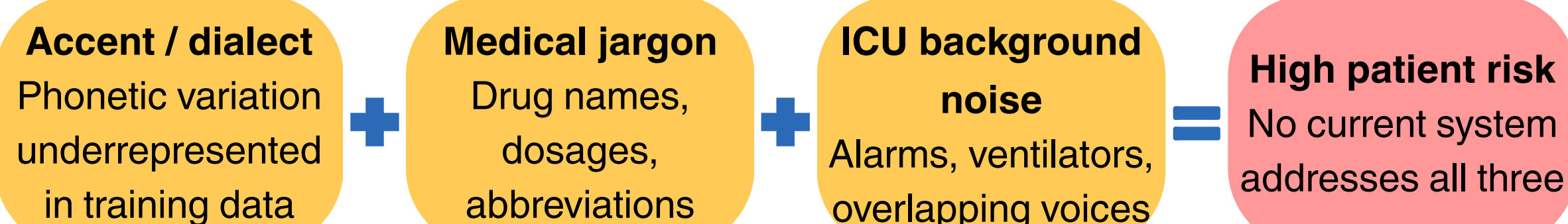
Source: Source: Mozilla Common Voice (Ardila et al., 2020)

How error rate grows with accent strength



Source: Illustrative trend based on del Rio et al. (2023) and Malisz et al. (2018) · Trend direction is literature-supported. Specific values are illustrative and should not be quoted as measured data.

#### Compounding error factors:



#### Clinical risk examples:

##### Dosage error

- "5 mg" → "50 mg" - 10x overdose risk

##### Drug name confusion

- "Noradrenalin" misrecognized or split

##### Silent failure

- Command not recognized, no alert raised

##### Real-time lag

- Complex models too slow for acute decisions

## Section 3

### The SASI solution

#### Four pillars

##### • Accent-robust speech recognition

Voice Conversion and synthetic data generation (GANs, phonemic augmentation) create a training set covering accents, dialects, and clinical shorthand.

##### • End-to-end speech transformer

A unified pipeline handles accent variation, ICU background noise (ventilators, alarms), and domain-specific language with Transfer Learning.

##### • Medical plausibility AI

A Safety Controlling System (SCS) checks every input against patient data, treatment plans, and clinical ontologies (SNOMED CT). Conflicts flagged before action.

##### • Interactive dialogue system

When confidence is low or a conflict detected, SASI proactively asks for clarification, keeping the clinician always in control, never silently failing.

